REMARKS

Claims 1-21 and 25-34 are pending in this application including independent claims 1, 16, 17, and 27 are independent. Claims 1, 14-17, 20, 27, and 33 have been amended. Support for the amendments to the claims can be found in specification and drawings as originally filed. No new matter has been added. Favorable reconsideration and allowance of the pending claims are requested.

Claim Rejections – 35 U.S.C. § 103(a)

Claims 1-13, 15, 17-19, 21, and 25-34

Claims 1-13, 15, 17-19, 21, and 25-34 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Borkar et al., "Automatic segmentation of text strings into structured records" ("Borkar") and in view of Ando et al., "Mostly-Unsupervised Statistical Segmentation of Japanese Sequences" ("Ando"). Applicants respectfully traverse this rejection.

When addressing independent claims 1, 17, and 27, the Office Action continues to rely on portions of Borkar that describe a Hidden Markov Model (HMM) that include states including start (S), state 1, state 2, state 3, and end (E). The Office Action also relies on portions of Borkar that describe restricting exploring paths that are invalid given the database of semantic relationships amongst symbols of different elements by modeling semantic constraints as a pair of symbol-state assignment that are invalid.

As admitted in the Office Action, Borkar fails to teach or suggest a state transition model based on an existing collection of data record that does not comprise manually segmented training data. When addressing such deficiencies of Borkar, the Office Action relies on portions of Ando related to a statistical method that utilizes unsegmented training data to sement Japanses kanji sequences.

While Applicants disagree with the grounds of rejection set forth in the Office Action, independent claims 1, 17, and 27 have been amended in order to advance prosecution. Applicants submit that the features recited by amended independent claim 1, 17, or 27 clearly distinguish over the teachings of Borkar and Ando, as well as the other references of record.

For instance, the HMM models described in Borkar clearly do not comprise a state transition model that categorizes tokens in database attribute values of the data records into positions based on a fixed beginning, middle, and trailing token topology that categorizes each boundary token of a database attribute value that includes multiple tokens into corresponding beginning and trailing positions and that categorizes each token that does not comprise a boundary token of a database attribute value into a middle position.

There is also no teaching or suggestion in Borkar of defining beginning, middle, and trailing state categories, wherein each state category includes states that accept tokens only if appearing in a corresponding one of said beginning, middle, and trailing positions. Further, Borkar fails to teach or suggest adjusting said states and probabilities associated with said states within said state categories in order to relax sequential specificity and account for erroneous token placement when evaluating tokens in the input string appearing in particular positions, wherein the state category corresponding to a particular position in which the token appears is adjusted to include states from another state category that accept tokens appearing in a different position.

Applicants respectfully submit that the DATAMOLD tool described in Borkar and relied upon in the present Office Action neither explicitly nor inherently discloses the features as recited by amended independent claims 1, 17, or 27. Applicants further submit that Ando also does not teach or suggest such features and does not remedy the deficiencies of Borkar with respect to such claims.

In view of the above, Applicants submit that none of the references, including Borkar and Ando, teaches or suggests all of the features recited by amended independent claims 1, 17, or 27. Consequently, even if Borkar and Ando could be combined, which Applicants do not admit, such combination still would not teach or suggest all of the features of amended independent claims 1, 17, or 27. Further, Applicants submit that there is no teaching, suggestion, or motivation to modify Borkar and/or Ando to include all of the recited features of amended independent claims 1, 17, or 27. Therefore, Applicants submit that Borkar and Ando, whether taken alone or in combination with each other, are insufficient to establish obviousness under § 103(a) with respect to amended independent claims 1, 17, or 27.

For at least the reasons set forth above, Applicants submit that amended independent claims 1, 17, and 27 are allowable and that dependent claims 2-13, 15, 18, 19, 21, 25, 26, and 28-34 are also allowable by virtue of their dependency from allowable claims, as well as on their own merits.

Accordingly, Applicants respectfully request reconsideration and withdrawal of the § 103(a) rejection of claims 1-13, 15, 17-19, 21, and 25-34.

Claims 14 and 20

Claims 14 and 20 stand rejected under 35 U.S.C. § 103(a) as being unpatentable over Borkar in view of Ando and further in view of United States Patent No. 5,095,432 to Reed ("Reed"). Applicants respectfully traverse this rejection.

As set forth above, Borkar and/or Ando do not teach or suggest all of the features of amended independent claims 1 or 17. Applicants further submit that the register vector grammar parsing algorithm described in Reed and relied upon in the Office Action does not remedy the deficiencies of Borkar and Ando with respect to the features recited by amended independent claims 1 or 17.

Consequently, even if Borkar, Ando, and/or Reed could be combined, which Applicants do not admit, such combination would not teach or suggest all of the features of amended independent claims 1 or 17. Further, Applicants submit that there is no teaching, suggestion, or motivation to modify Borkar, Ando and/or Reed to include all of the recited features of amended independent claims 1 or 17. Therefore, Applicants submit that Borkar, Ando and Reed, whether taken alone or in combination with each other, are insufficient to establish obviousness under § 103(a) with respect to amended independent claims 1 or 17.

For at least the reasons set forth above, Applicants submit that amended independent claims 1 and 17 are allowable and that dependent claims 14 and 20 are also allowable by virtue of their dependency from allowable claims, as well as on their own merits.

Accordingly, Applicants respectfully request reconsideration and withdrawal of the § 103(a) rejection of claims 14 and 20.

Claim 16

Claim 16 stands rejected under 35 U.S.C. § 103(a) as being unpatentable over Borkar in view of Ando and further in view of United States Published Patent Application 2006/0235811 to Fairweather ("Fairweather"). Applicants respectfully traverse this rejection.

While Applicants disagree with the grounds of rejection set forth in the Office Action, independent claim 16 has been amended in order to advance prosecution. Applicants submit that the features recited by amended independent claim 16 clearly distinguish over the teachings of Borkar, Ando, and Fairweather, as well as the other references of record.

For instance Borkar does not disclose a state model that categorizes substrings within database attribute values into positions based on a fixed beginning, middle, and trailing token topology for said attribute that categorizes beginning and trailing substrings of a database attribute value that includes multiple substrings into corresponding beginning and trailing positions, categorizes each substring of the database attribute value that does not comprise a beginning or trailing substring into a middle position, andaccepts a null token for an empty attribute component.

There is also no teaching or suggestion in Borkar of defining beginning, middle, and trailing state categories, wherein each state category includes states that accept tokens only if appearing in a corresponding one of said beginning, middle, and trailing positions. Further, Borkar fails to teach or suggest adjusting said states and probabilities associated with said states within said state categories in order to relax sequential specificity and account for erroneous token placement when evaluating tokens in the input string appearing in particular positions, wherein the state category corresponding to a particular position in which the token appears is adjusted to include states from another state category that accept tokens appearing in a different position.

Applicants respectfully submit that the DATAMOLD tool described in Borkar and relied upon in the present Office Action neither explicitly nor inherently discloses the features recited by amended independent claim 16. Applicants further submit that Ando and Fairweather also do not teach or suggest such features and do not remedy the deficiencies of Borkar with respect to such claims.

In view of the above, Applicants submit that none of the references, including Borkar, Ando, and Fairweather, teaches or suggests all of the features recited by amended independent claim 16. Consequently, even if Borkar, Ando, and/or Fairweather could be combined, which Applicants do not admit, such combination would not teach or suggest all of the features of amended independent claim 16. Further, Applicants submit that there is no teaching, suggestion, or motivation to modify Borkar, Ando, and/or Fairweather to include all of the recited features of amended independent claim 16. Therefore, Applicants submit that Borkar, Ando, and Fairweather, whether taken alone or in combination with each other, are insufficient to establish obviousness under § 103(a) with respect to amended independent claim 16.

For at least the reasons set forth above, Applicants submit that amended independent claim 16 is allowable.

Accordingly, Applicants respectfully request reconsideration and withdrawal of the § 103(a) rejection of claim 16.

Conclusion

It is believed that claims 1-21 and 25-34 are in condition for allowance. Accordingly, a timely Notice of Allowance to this effect is earnestly solicited.

Applicants do not otherwise concede, however, the correctness of the Office Action with respect to any of the limitations of the independent claims and dependent claims discussed above. Accordingly, Applicants hereby reserve the right to make additional arguments as may be necessary to further distinguish the claims from the cited references, taken alone or in combination, based on additional features contained in the independent or dependent claims that were not discussed above. A detailed discussion of these differences is believed to be unnecessary at this time in view of the basic differences in the independent claims pointed out above.

The Examiner is invited to contact the undersigned to discuss any matter concerning this application.

The Office is hereby authorized to charge any additional fees or credit any overpayments under 37 C.F.R. § 1.16 or § 1.17 deposit account 50-0463.

Respectfully submitted,

/Robert V. Racunas/

Robert V. Racunas, Reg. No. 43,027
Under 37 CFR 1.34(a)

Dated: July 17, 2009
Microsoft Corporation
One Microsoft Way
Redmond WA 98052-6399
Direct Telephone (425) 707-9382